

REMARKS/ARGUMENTS

Claims 1-8, 10-22, 24-35 and 37-43 are pending in the present application. Claims 1, 11, 15, 25, 29, 38 and 41-43 have been amended, and Claims 9, 23 and 36 have been cancelled, herewith. Reconsideration of the claims is respectfully requested.

I. 35 U.S.C. § 101

The Examiner rejected Claims 1-43 under 35 U.S.C. § 101 as being directed towards non-statutory subject matter. This rejection is respectfully traversed.

In rejecting all of Claims 1-43, the Examiner states that such claims do not produce a useful, concrete and tangible result. Applicants urge that the claimed invention as a whole is useful and accomplishes a practical application. For example, with respect to Claim 1 (and dependent Claims 2-14), such claim is directed to a method for *selecting data sets that are to be used* by a predictive algorithm. The selection of the data sets is based upon data network geographic information. A statistical distribution of a training data set is generated, such statistical distribution being a tangible result. A statistical distribution of a testing data set is generated, such statistical distribution being a tangible result. These two tangible results – the generated statistical distribution of the training data set and the generated statistical distribution of the testing data set are compared to identify a discrepancy of these generated statistical distributions with respect data network geographic information. This identified discrepancy is also a tangible result, and is *used to modify the selection of entries* in either or both the training data set and the testing data set. This modified selection of entries is also a tangible result, and is subsequently used by the predictive algorithm such that the predictive algorithm takes into account the data network geographic information by using the modified selection of entries. In summary, Claim 1 is not merely directed to a predictive algorithm, but rather is directed to a particular technique for *selecting parameters used by such algorithm*. Therefore, Claim 1 does produce a useful, concrete and tangible result of *generating two distributions* of two data sets, *identifying a discrepancy* of such generated distributions with respect to data network geographic information, and then *modifying the selection of entries* for one or both of the training and testing data set which is then used by a predictive algorithm to thereby enhance the predictions of such predictive algorithm by taking into account data network geographic information as represented in the training and testing data sets.

Further with respect to Claim 7, such claim recites additional useful, concrete and tangible results in that it recites generation of *recommendations* for improving selection of entries for either or both the training and testing data sets. Such recommendations are then advantageously used to re-generate one or both of the statistical distributions.

Further with respect to Claim 8, such claim recites additional useful, concrete, and tangible results, as the training data set and the testing data set are selected from a customer information database that comprises information with respect to customers who have purchased goods and services over a data network, and the data network geographic information pertains to such geographic information of the data network.

Further with respect to Claim 11, such claim recites additional useful, concrete and tangible results – specifically, the generation of both (i) a composite data set and (ii) a composite distribution of such generated composite data set.

Further with respect to Claim 12, such claim recites additional useful, concrete and tangible results in that such claim recites a step of training the predictive algorithm and thus the result is a trained predictive algorithm.

Claim 15 (and dependent Claims 16-28) is statutory for similar reasons to those given above with respect to Claim 1, and in addition such claim is specifically directed to an apparatus that comprises (i) a statistical engine, (ii) a comparison engine coupled to the statistical engine, and (iii) a predictive algorithm device which uses both (a) the modified selection of entries and (b) the predictive algorithm.

Claim 21 is further shown to be statutory for similar reasons to those further reasons given above with respect to Claim 7.

Claim 22 is further shown to be statutory for similar reasons to those further reasons given above with respect to Claim 8.

Claim 25 is further shown to be statutory for similar reasons to those further reasons given above with respect to Claim 11.

Claim 26 is further shown to be statutory for similar reasons to those further reasons given above with respect to Claim 12.

Claim 29 (and dependent Claims 30-40) is statutory for similar reasons to those given above with respect to Claim 1.

Claim 35 is further shown to be statutory for similar reasons to those further reasons given above with respect to Claim 7.

Claim 38 is further shown to be statutory for similar reasons to those further reasons given above with respect to Claim 11.

Claim 39 is further shown to be statutory for similar reasons to those further reasons given above with respect to Claim 12.

With respect to Claim 41 (and similarly for Claims 42 and 43), the useful, concrete and tangible result is the prediction of customer behavior based on the data network geographic information – i.e. predicted customer behavior. In addition, another useful, concrete and tangible result of Claim 41 is data

network geographic information regarding a plurality of customers, which is (i) obtained and (ii) used to train a predictive algorithm.

Therefore, the rejection of Claims 1-43 under 35 U.S.C. § 101 has been overcome.

II. 35 U.S.C. § 112, First Paragraph

The Examiner objected to the Specification under 35 U.S.C. § 112, first paragraph, as failing to adequately teach how to make and/or use the invention in Claims 1, 15, 29 and 41-43. Additionally, the Examiner rejected the claims under the same reasons. This rejection is respectfully traversed.

As to Claims 1, 15, 29 and 41-43, the Examiner states that nowhere in the Specification is it explained how the predictive algorithm would predict customer behavior based upon network geographic location. Applicants will now show where this is explained in the Specification.

As shown in Figure 4, a set of customers 400 for which information has been obtained are present in a data network geographical area. These customers 400 are geographically located in the data network in clusters due to their affiliation with other customers that navigate the data network in a similar manner or are geographically located in the data network near other customers. For example, customers that navigate the data network using similar type search terms may be required to traverse the same number, or close to the same number, of links in order to arrive at a destination web site or web page. Because of this, these customers may be geographically located close to one another in the data network since it requires the same amount of travel distance for these customers to arrive at other data network web sites. From these customers 400 a customer database is generated 410 (Specification page 19, line 19 – page 20, line 8). From the customer database 410, a set of training data 420 and testing data 430 are generated. In known systems, these sets of data 420 and 430 are generated using a random selection process. Based on this random selection process, various ones of the customers in the customer database 410 are selected for inclusion into the training data set 420 and the testing data set 430. As can be seen from Figure 4, by selecting customers randomly from the customer database 410, it is possible that some of the clusters of customers may not be represented in the training and testing data sets 420 and 430. Moreover, the training data set 420 and the testing data set 430 may not be commonly representative of the same clusters of customers. In other words, the training data set 420 may contain customers from clusters 1 and 3 while the testing data set 430 may contain customers selected from clusters 1 and 2. Because of the discrepancies between the training and testing data sets 420 and 430 with the customer database 410, certain types of customers may be over-represented and other types of customers may be under-represented. As a result, the predictive algorithm may not accurately represent the behavior of potential customers. Moreover because of the discrepancies between the training and testing data sets 420 and 430, the predictive algorithm may be trained improperly. That is, the training data set 420 may train

the predictive algorithm to output a particular predicted customer behavior based on a particular input. However, the testing data set 430 may indicate a different customer behavior based on the same input due to the differences in the customer clusters represented in the training data set 420 and the testing data set 430 (Specification page 20, line 19 – page 21, line 27).

For example, as shown in Figure 4, the training data set 420 is predominately comprised of customers from clusters A, B and C. Assume that customers in clusters A and B are very good customer candidates for new electronic items while customers in group C are only mildly good customer candidates for new electronic items. Based on this training data, if a commercial web site at data network location X were interested in introducing a new electronic item, the predictive algorithm may indicate that there is a high likelihood of customer demand for the new electronic item from customers in clusters A and B. However, in actuality, assume that customers in clusters A and B are less likely to navigate the data network from their data network location to the data network location X due to the amount of interaction required, i.e. the size of the user click stream. Thus, the predictive algorithm will provide an erroneous result. Moreover, if the testing data contains customers from clusters A, B, D and E, the customer behaviors in the testing data will be different from that of customers in the training data set (comprising clusters A, B and C). As a result, the testing data and the training data are not consistent and erroneous customer behavior predictions will arise. Thus, data network geographic effects of clustering must be taken into account when selecting customers to be included in training and testing data sets of a customer behavior predictive algorithm (Specification page 22, line 1 – page 23, line 3).

With the present invention, the discrepancies between a testing data set and a training data set are identified. Furthermore, the discrepancies between both the testing data set and the training data set and the customer database are identified. The discrepancies are identified based on a data network geographical characteristic such as a number of links or the size of a user click stream. The normalized frequency distributions of the number of links and/or user click stream in the training data set are compared to the normalized frequency distributions of the testing data set. If the differences between the frequency distributions is above a predetermined tolerance, the two data sets are too different to provide accurate training of the predictive algorithm when taking data network geographical influences into account. This same procedure may be performed with regard to the frequency distribution of the customer database (Specification page 23, lines 4 – 21).

In order to compare the frequency distributions, the mean, mode and/or standard deviations of the frequency distributions may be compared with one another to determine if the frequency distributions are similar within a predetermined tolerance. The mean is a representation of the average of the frequency distribution. The mode is a representation of the most frequently occurring value in the data set. The standard deviation is a measure of dispersion in a set of data. Based on these quantities for each

frequency distribution, a comparison of the frequency distributions may be made to determine if they adequately represent the customer population clusters in the customer database. *If they do not, the present invention may, based on the relative discrepancies of the various data sets, make recommendations as to how to better select training and testing data sets that represent the data network geographic clustering of customers.* For example, if the relative discrepancy between a testing data set and a training data set are such that the training data set does not contain enough customers to represent all of the necessary clusters in the testing data set, the training data set may need to be increased in size. Similarly, if the testing data set and/or training data set do not contain enough customers to represent all of the clusters in the customer database, the testing and training data sets may need to be increased. In such cases, the same random selection algorithm may be used and the same seed value of the random selection algorithm may be used with the number of customers selected being increased. Moreover, the testing data set and training data sets may be combined to form a composite data set which may be compared to the customer database. In combining the two data sets, customers appearing in a first data set, and not in the second data set, are added to the composite data set, and vice versa (Specification page 23, line 23 – page 25, line 4).

The frequency distribution of the composite data set may be compared to the frequency distribution of the customer database, in the manner described above, to determine if the composite represents the customer clusters appropriately. If the composite data set does represent the customer clusters of the customer database appropriately, the composite data set may be used to train the predictive algorithm. If the composite data set does not represent the customer clusters of the customer database appropriately, a new random selection algorithm may need to be used or a new seed value of a random selection algorithm may need to be used. *In this way, the selection of training and testing data is modified such that the training and testing data better represents actual customer behavior based on data network geographical influences* (Specification page 25, lines 5-20).

Figure 6 is a flowchart outlining an exemplary operation of the present invention. As shown in Figure 6, the operation starts with gathering customer database information (step 610). The customer database information is then used as a basis for selecting a training data set and/or testing data set (step 620). Frequency distribution information of a number of data network links and/or user click stream to a web site of interest is calculated for each of the training data set, testing data set and customer database data set (step 630). The frequency distribution information for each of these data sets is compared and evaluated to determine if differences exceed a predetermined tolerance (step 640). *A determination is made as to whether differences in the frequency distribution information is beyond a predetermined tolerance (step 650). If so, recommendations are generated based on the particular differences (step 660) and the operation returns to step 620 where the training and testing data sets are again determined in*

view of the recommendations. If the differences in frequency distribution information are not beyond the predetermined tolerance, the training data set and testing data set are used to train the predictive algorithm (step 670) and the operation ends. Thereafter, the predictive algorithm may be used to generate customer behavior predictions *taking into account the data network geographical influences of customers as represented in the training and testing data sets* (page 47, line 21 – page 48, line 22).

Therefore, the objection of the Specification and rejection of Claims 1, 15, 29 and 41-43 under 35 U.S.C. § 112, first paragraph has been overcome, as the Specification does in fact describe in detail *how parameters used by the predictive algorithm are modified to improve predicted customer behavior based upon network geographic location.*

III. 35 U.S.C. § 103, Obviousness

The Examiner rejected Claims 1-40 under 35 U.S.C. § 103 as being unpatentable over Menon et al (U.S. 5,537,488) in view of Wu (U.S. 6,741,967) and further in view of Applicant's background of the Invention. This rejection is respectfully traversed.

The present invention of Claim 1 is directed to an improved technique for selecting data sets for use with a predictive algorithm. A statistical distribution of a training data set is compared with a statistical distribution of a testing data set to identify a discrepancy between these distributions with respect to data network geographic information. *Based upon such comparison and its associated discrepancy identification, selection of entries in the training data set and/or testing data set is modified.* These modified entries are then used by the predictive algorithm, thereby taking into account the influences of data network geography when using the predictive algorithm. None of the cited references makes any mention of *using data network geographic information to modify entries of the testing or training data sets* that is used by a predictive algorithm. The Examiner characterizes Wu's teachings as a system that determines customer's navigational path through web sites or web pages and predicts if an increase in a customer's purchase rate was the result of an improvement in the navigational path (citing Wu column 36, lines 24-30). Applicants respectfully submit that Wu does not teach any type of customer navigational path determination, as the reference merely alludes to *navigational cues* that exist in a given web site. Navigational cues are just that, cues or hints that are used to inform a user of certain information. There is no teaching or suggestion of any type of determination with respect to actual network geographic information being used, either to modify entries of testing/training data sets (as claimed), or any other type of use. By analogy, a highway may have signs posted directing drivers to take a certain route, or stating what the maximum speed limit is. Such posting of signs, akin to navigational cues placed on a web site per Wu, provide no indication as to whether a driver actually followed the suggested route or abided by the posted speed limit. They are merely passive postings of information.

Similarly, Wu's teaching of determining whether an increase in purchase rate was *based on improved navigational cues* (i.e. the posting of information on a web site) provides no information on the navigational path *actually taken by a user*. Thus, this cited reference does not teach or otherwise suggest *using data network geographic information to modify entries of the testing or training data sets* that is used by a predictive algorithm, as per Claim 1. It is therefore respectfully submitted that the Examiner has failed to properly establish a prima facie showing of obviousness with respect to Claim 1¹. Accordingly, the burden has not shifted to Applicants to overcome an obviousness assertion². In addition, as a proper prima facie showing of obviousness has not been established, Claim 1 has been improperly rejected³.

In any event, in order to expeditiously move this case to issuance, Applicants have amended Claim 1 to include features of Claim 9 (which is thus being cancelled herewith). Applicants further traverse the rejection of amended Claim 1 for further reasons given below in discussing the rejection of Claim 9.

With respect to Claims 2, 3, 5, 6, 10, 16, 17, 19, 20, 24, 30, 31, 33, 34 and 37, the Examiner acknowledges that the cited references do not teach the claimed features recited in such claims, but states:

"However, the same argument made to claim 1 regarding this missing limitation is also made in claim 2."

Applicants urge that mere reliance on the reasoning given in rejecting Claim 1 as the sole basis for the reasoning in rejecting Claims 2, 3, 5, 6, 10, 16, 17, 19, 20, 24, 30, 31, 33, 34 and 37 is clearly erroneous, as these additional claims recite additional features not recited in Claim 1, and thus the Examiner has not explicitly addressed how the specific features recited in Claims 2, 3, 5, 6, 10, 16, 17, 19, 20, 24, 30, 31, 33, 34 and 37 are taught or suggested by the cited references. Therefore, the Examiner has failed to establish a prima facie showing of obviousness with respect to Claims 2, 3, 5, 6, 10, 16, 17, 19,

¹ To establish prima facie obviousness of a claimed invention, all of the claim limitations must be taught or suggested by the prior art (emphasis added by Applicants). MPEP 2143.03. See also, *In re Royka*, 490 F.2d 580 (C.C.P.A. 1974).

² In rejecting claims under 35 U.S.C. Section 103, the examiner bears the initial burden of presenting a prima facie case of obviousness. *In re Oeriker*, 977 F.2d 1443, 1445, 24 USPQ2d 1443, 1444 (Fed. Cir. 1992). Only if that burden is met, does the burden of coming forward with evidence or argument shift to the applicant. *Id.*

³ If the examiner fails to establish a prima facie case, the rejection is improper and will be overturned. *In re Fine*, 837 F.2d 1071, 1074, 5 USPQ2d 1596, 1598 (Fed. Cir. 1988).

20, 24, 30, 31, 33, 34 and 37, and accordingly the burden has not shifted to Applicants to rebut such improper obviousness assertion. In addition, as a proper *prima facie* case of obviousness has not been established by the Examiner, the rejection of Claims 2, 3, 5, 6, 10, 16, 17, 19, 20, 24, 30, 31, 33, 34 and 37 is improper.

Further with respect to Claim 7, it is urged that none of the cited references teach or suggest the claimed feature of *generating recommendations for improving selection of entries* in either or both of the training and testing data set. The passage cited by the Examiner in rejecting Claim 7 merely states that a new category is defined (if the correlation is below a threshold). Such 'definition of a new category' does not provide any type of *recommendation for improving selection of entries*, as expressly recited in Claim 7, and thus Claim 7 is further shown to have been erroneously rejected as there are additional claimed features not taught or suggested by any of the cited references.

Further with respect to Claim 9 (whose features are now a part of Claim 1), Applicants urge that none of the cited references teach (or otherwise suggest) the claimed step of "comparing at least one of the first statistical distribution and the second distribution to a distribution of a customer database". As can be seen, this claim is directed to comparing one or more of the first and second distributions *with another distribution – the distribution of a customer database*. The cited Menon reference does not teach (or otherwise suggest) a *distribution* of a customer database, and hence it necessarily follows that it does not teach (or otherwise suggest) any comparing step being made with such (missing) distribution of a customer database. In rejecting Claim 9, the Examiner cites Menon col. 6, line 57 – column 7, line 21 as teaching the features of Claim 9. Applicants urge that such passage describes details of how to group training patterns into categories in order to generate a training histogram, as claimed by Menon in Claim 24 (col. 20, lines 55-60). This passage deals with training patterns and the labeling of these training patterns' associated categories. The calculations described are *only with respect to training patterns* – albeit organized into different groups or categories. Importantly, there is no teaching (or suggestion) of comparing such training patterns to a *distribution of a customer database*, as expressly recited in Claim 9. Thus, it is further shown that a *prima facie* case of obviousness has not been established with respect to Claim 9.

Applicants traverse the rejection of Claims 15-40 for similar reasons to those given above. Therefore, the rejection of Claims 1-40 under 35 U.S.C. § 103 has been overcome.

IV. 35 U.S.C. § 103. Obviousness

The Examiner rejected Claims 41-43 under 35 U.S.C. § 103 as being unpatentable over Wu (U.S. 6,741,967) in view of Applicant's background of the Invention. This rejection is respectfully traversed.

Applicants have amended such claims to clearly differentiate the claimed invention recited therein from the teachings of the cited references. As described above with respect to Claim 1, Wu merely teaches determining if a purchase rate increase is attributable to better *navigational cues* – i.e. information given to a user, whereas Claim 41 is directed to *using actual customer geographic information to train a predictive algorithm*. None of the cited references teach or otherwise suggest using such customer geographic information to train a predictive algorithm, and thus the amendment to Claim 41 has overcome the present 35 USC 103 rejection.

Applicants traverse the rejection of Claims 42 and 43 for similar reasons to those given above with respect to Claim 41.

Therefore, the rejection of Claims 41-43 under 35 U.S.C. § 103 has been overcome.

V. Conclusion

It is respectfully urged that the subject application is patentable over the cited references and is now in condition for allowance. The Examiner is invited to call the undersigned at the below-listed telephone number if in the opinion of the Examiner such a telephone conference would expedite or aid the prosecution and examination of this application.

DATE: June 12, 2006

Respectfully submitted,



Brian D. Owens
Reg. No. 55,517
Wayne P. Bailey
Reg. No. 34,289
Yee & Associates, P.C.
P.O. Box 802333
Dallas, TX 75380
(972) 385-8777
Attorneys for Applicant